KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and Applied Economics

Department of Economics

Community Targeting for Poverty Reduction in Burkina Faso.

by

David BIGMAN
Stefan DERCON
Dominique GUILLAUME
Michel LAMBOTTE

Development Economics

Center for Economic Studies
Discussions Paper Series (DPS) 99.10
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

March 1999

DISCUSSION PAPER

Community Targeting for Poverty Reduction in Burkina Faso

by*

David Bigman

Stefan Dercon

Dominique Guillaume

Michel Lambotte

Abstract: The paper develops a method for targeting anti-poverty programs and public projects on poor communities in rural and urban areas. The method is based on the application of an extensive data-set from a large number of sources and the integration of the entire data-set in a Geographical Information System. This data-set includes data from the population census, household-level data from a variety of surveys, community-level data on the local road infrastructure, public facilities, water points, etc., and department-level data on the agro-climatic conditions. An econometric model that estimates the impact of household-, community-, and department-level variables on households' consumption has been used to identify the key explanatory variables that determine the standard of living in rural and urban areas. This model was applied to predict poverty indicators for 3871 rural and urban communities across the country and to provide a mapping of the spatial distribution of poverty in Burkina Faso. Simulation analysis was subsequently conducted to assess the effectiveness of village-level targeting based on these predictions of the poverty indicators. The results show that village-level targeting based on these predictions provides an improvement over regional targeting by reducing the leakage of the targeted program and the percentage of the population that remains undercovered.

This study was part of a project on Geographical Targeting in Burkina Faso. We would like to thank the Belgian Trust Fund for providing financial support for this study. We would also like to thank the following institutions in Burkina Faso that assisted us in this work and provided us with the necessary data: IGB, DGH/DRH, DSAP, ONEA, INSD, DMN, MEBA, DEP-SANTE, DMP, FEWS. We are grateful to Bart Capéau, Dirk Van de Gaer, Jesko Hentschel, Peter Lanjouw, Pramila Krishnan, three anonymous referees and participants in the World Bank workshop on Geographical Targeting for Poverty Reduction and Rural Development for useful comments.

---

* David Bigman was at the World Bank while this work was done and is currently at ISNAR; Stefan Dercon and Dominique Guillaume are at the Catholic University of Leuven, Belgium and at Oxford University - Centre for the Study of African Economies; and Michel Lambotte is at I-Mage, Mons, Belgium.

1. Introduction

The budgetary and social pressures to increase the impact of health, education and rural development projects and programs on the poor gave strong impetus to improve the targeting of public projects and programs. Undifferentiated transfers that cover the entire population, such as general food subsidies, proved to be beyond the budget constraints of most developing countries, and their benefits go disproportionately to the non-poor[1]. In many developing countries, particularly in SSA, targeting criteria that cover large geographical areas or large population groups are also likely to be too costly and too ineffective; a program that is targeted on the entire rural population, for example, will cover in these countries the majority of the country's total population — poor and non-poor.

More accurate targeting requires the choice of criteria that can effectively identify the eligible recipients. Such criteria can be narrowly defined — at the level of individual households, or they can be more broad-based — e.g., at the level of the region or the province — by identifying the geographical areas or the population groups that have a higher than average incidence of poverty (Van de Walle, 1991). Narrow targeting at the household level is very information-intensive, and the necessary information is very costly. Identification of the eligible households requires complex and expensive means-testing, and even in many *developed* countries it is only partly successful — despite the wide range of data that is available in these countries on individual households — and a large portion of the benefits "leaks" to non-eligible households. In most *developing* countries, the information on individual households — particularly the poor households — which is necessary for means-testing is not available, and the scope for narrow targeting at the household level is therefore very limited. As an alternative to direct means-testing, the standard household income and expenditure surveys, such as the World Bank's Living Standard Measurement Surveys (LSMS), can be used to identify the more general characteristics of the poor and thereby determine a set of indicators, such as the number of children or the place of residence, that can distinguish the poor and thereby establish eligibility without resorting to direct means testing. Using LSMS data for Côte d'Ivoire, Glewwe (1991) examined the trade-off between the use of a refined and exhaustive set of indicators for narrow targeting and the costs of collecting the information on these indicators and concluded that, in the case of Côte d'Ivoire, a rather limited set of community and household indicators proved to be quite effective in identifying the poor households. However, the incentives for households to change or lie about their characteristics in order to qualify for the program once these indicators are determined as eligibility criteria, can significantly reduce its effectiveness and blow up the budgetary costs.[2] This, together with the high costs of administering a program at the household level and the peril that these eligibility criteria will leave out many of the country's poor, deterred the governments of most developing countries from targeting social welfare programs on individual households.

Geographical targeting at the level of the province or the region may offer an effective approach for reaching the poor in countries where there are substantial disparities in living conditions between geographical areas, and where administering these programs is relatively less complex because the local administration is already in place. In India, the allocation of central government disbursements across states has long been determined, in part, by the large disparities between states in their levels of poverty. The decision to locate rural development projects in backward regions has become the center of India's poverty oriented agricultural development strategy. However, even in countries where the poor concentrate in certain states or regions, geographical targeting at the level of large administrative areas is likely to entail considerable leakage of benefits to the non-poor that live in the target areas, while failing to cover the poor that live in other areas. Although targeting at these high levels of geographical aggregation is likely to be more effective in reducing leakage and enhancing coverage than general, non-targeted programs, quantitatively the effectiveness of these programs tend to be rather small (Ravallion, 1993; Baker and Grosh, 1994; Ravallion, 1996).[3] Ravallion (1993) evaluated the costs and effects of geographical targeting at the level of the province in Indonesia and concluded that, although this program offer clear gains in terms of poverty alleviation, the magnitude of these gains is rather small. Baker and Grosh (1994) analyzed geographical targeting in Venezuela, Mexico and Jamaica and concluded that targeting priority regions can be an effective mechanism of transferring benefits to the poor, but, with a given budget constraint, poverty reduction is greater the more finely

[1] For simulated examples from Latin America, see Baker and Grosh (1994).
[2] See also Besley and Kanbur (1991).
[3] These programs may also provide incentives to households to move to the targeted areas, thereby defeating the purpose of the program and raising its costs.

defined and the narrower the target areas are, and the greatest reduction in poverty is achieved when the target areas are municipalities or villages.

Narrow geographical targeting at the level of the village or the urban community can reduce the leakage of benefits to the non-poor in countries or regions where the socio-economic conditions and the standard of living of the majority of the population in the villages or the urban communities are rather similar. Often in these countries and regions, many of the households in a village have rather similar sources of income and all households are affected by the same agro-climatic and geographic conditions — including the road condition, the distance to the nearest town, and the availability of public facilities for health, education, water supply, etc. Consequently, income inequality between individuals in these countries or regions is often due, to a considerable degree, to income differences *between* villages and only to a lesser degree to income differences between individuals *within* villages.

Targeting at the lower geographical level of the district or the village requires, however, much more information on the spatial distribution of poverty across districts or villages and on the characteristics of the poor population in these areas. However, the information on the standard of living of the population is provided, in most developing countries, by a household survey, and the size of the sample in the standard survey is far too small to allow an estimation of the incidence of poverty at the level of the village or the district for the entire country. At the present, the LSMS surveys can provide a map of the spatial distribution of poverty only between the country's main regions. Some countries that resort to geographical targeting use an alternative set of indicators to estimate the geographic distribution of poverty and establish criteria for targeting that is based on more readily available indicators such as access to public services, the percentage of the school-age children that attend school, the prevalence of certain illnesses that are associated with malnutrition, etc. All too often, however, these indicators are not sufficiently correlated with the welfare indicators of the local population, and this may lead to targeting errors in determining eligibility and the ineffective use of resources (Hentschel et al., 1998).

The objective of this paper is to present a method for narrow geographical targeting at the level of rural villages and urban communities. The method is based on the construction of a very large data-set from a wide variety of sources in the form of a

Geographic Information System (GIS) and the use of these data to provide a mapping of poverty at the community and the province levels. This data-set includes several strata of information: first, demographic and socio-economic information at the household level from a variety of surveys; second, village- and community-level information, including demographic information from the population census, the distance to the urban centers, the condition of the road infrastructure, the availability and quality of public services, the sources of drinking water, etc.; third, department- or region-level information on agro-climatic and geographic conditions, including the location of the main towns and main transport routes. The entire data-set was integrated at the level of the village or the urban community using geo-referencing, and organized in the form of a GIS database. The second step is to use this data-set in an econometric analysis that uses also the detailed data of a household survey on households' to construct a prediction model of — using household-, community-, and department-level variables. These variables were selected, however, from the GIS database, and they include therefore only variables for which mean values were available for *all* communities in the country. The third step is to apply the predictions of this model in order to derive estimates of the *average* level of well-being of the households in a community and the incidence of poverty in that community for *all* the communities in the country. These estimates were derived on the basis of the community and Department data that were available for *all* communities. These estimates determined, in turn, the spatial distribution of poverty in the country at the village level.[4] This method has been applied for Burkina-Faso, using the relatively detailed household data of the Priority Survey (PS).

The plan of the Paper is as follows: Section 2 describes the method and the econometric model that were used to estimate households' well-being from the sample of the PS and the method of applying these estimates in order to predict poverty levels at the community level in the communities outside the sample of the PS. Section 3 provides the details of the different data sources and presents their organization in a Geographical Information System. Section 4 discusses the specification of the prediction model. Section 5 presents the results of the econometric analysis of the household survey. Section 6 demonstrates the application of these estimates in order to predict the incidence of poverty in the villages outside the sample of the PS. Section 7 presents some

[4] Due to data limitations discussed below, the complete data set necessary for the predictions was available only for 3871 out of the country's 6821 villages.

simulations on the effectiveness of the targeting system. Section 8 offers some concluding remarks.

2. Methodology

The econometric analysis in this study has two parts: In the first part, a prediction model for household consumption is estimated, using the household data of the PS, and the community data from all other sources, in order to determine the variables that best explain households' consumption levels and households' poverty. The explanatory variables in that model are selected so that only variables for which we had data for all villages outside the PS sample are included. In the second part, the prediction model was used to determine the levels of welfare at the village-level for all the villages outside the sample of the PS, using the village-level data of the explanatory variables from the GIS database. In line with similar studies on this subject, we use consumption per 'standard adult' ('adult equivalent') as our welfare indicator at the household level and focus on the poverty incidence, measured by the Headcount index, as the measure of poverty. [5]

Let c_{ij} denote the level of consumption per standard adult in household *i*, residing in community *j*. Let *z* denote the *poverty line* and let $y_{ij} = c_{ij}/z$ be the normalized welfare indicator per standard adult. The analysis will be conducted in terms of the natural logarithms of y_{ij}. [6] The Headcount index H_j, which measures the relative size of the poor population in community *j*, is equal to the mean value of the individual poverty indicators H_{ij} — which indicate the probability that the household *ij* is poor — over all the individuals in that community. The individual poverty indicator is determined by the normalized welfare function as follows:

$$H_{ij} = 1 \quad \text{if } \ln y_{ij} < 0 \qquad (1)$$

$$H_{ij} = 0 \quad \text{if } \ln y_{ij} \geq 0$$

In the construction of the prediction model, the individual welfare indicator is modeled as a function of a vector of household and community explanatory variables X_{ij} and a residual term u_{ij}, which is assumed to be normally distributed with $u_{ij} \sim N(0,\sigma_j^2)$ — thereby allowing for village level heteroscedasticity. The prediction model is thus given by:

[5] Simple nutritional adult equivalent scales were used, using 0.7 for a child for 5-15 years and 0.3 for a younger child, and each adult counted as one.
[6] For a poor person, therefore: $y_{ij} < 1$, or $\ln y_{ij} < 0$.

$$\ln y_{ij} = \beta' X_{ij} + u_{ij} \quad (2)$$

As noted earlier, the explanatory variables were selected only if their *mean* values were available for *all* villages in the GIS database. They include community characteristics as well as mean values of household characteristics for all households in the community such as average household composition, average literacy rates, etc. Equation (2) can be estimated by means of maximum likelihood, with $u_{ij} \sim N(0, \sigma^2.\exp(\gamma' X_j^V)$), where X_j^V is the mean values of the explanatory variables in community *j*, in order to correct for the heteroscedasticity, and obtain the estimators *b* and s_j, of the parameters β and σ_j. These estimators and the set of explanatory variables can be used to predict the community's mean consumption for all communities outside the PS sample. Mean consumption in a community is, however, not necessarily a good predictor of poverty, since the poverty measure is a function of not only mean consumption, but also of the *distribution* of consumption in the community. The term s_j represents one part of that distribution, since the *within-community* variance is the sum of the variance of the regression *and* the deviation of the predicted household level consumption from the predicted mean level of consumption. Both components are therefore part of the overall measure of the within-village distribution of consumption.[7]

Using these estimators and the set of explanatory variables, a consistent estimate of the probability that household *ij* with characteristics X_{ij} is poor can then be expressed as:

$$E(H_{ij} \mid X_{ij}, b, s_j) = \text{Prob}\,(u_i < -b'X_{ij}) = \Phi(-b'X_{ij}/s_j) \quad (3)$$

where $\Phi(.)$ is the cumulative normal distribution. The predicted level of the incidence of poverty *per community j* is determined, from this equation, as:

$$E(H_j \mid X_{ij}, b, s_j) = E\,(\Phi\,(-b'X_{ij}/s_j)) \quad (4)$$

If complete information on the variables X_{ij} was available for *all* households and *all* villages in the country, this prediction would have been fairly straightforward: equation (3) would then be used to estimate the probability that each of the households in the village is poor, and equation (4) would be used to predict the incidence of poverty in the community -- across all

[7] Note that the within-village variance of consumption can be written as:
$E[(Y_{ij}-E(Y_j))^2] = E[(b'X_{ij}-b' X_j^V)^2] + s_j^2$,
in which Y_j is the mean level of consumption in the village. In words, the variance of consumption is the sum of the squared deviation of predicted household consumption from predicted mean consumption per

villages outside the sample.[8] However, in Burkina Faso the only data available for all the villages outside the PS sample are the *mean* values of the explanatory variables X_j^V per community. Since (4) is non-linear, these variables cannot be simply used to predict the village-level poverty incidence. Using Taylor-expansions, it is nevertheless possible to obtain an approximation. For this purpose, (4) can be expanded around (- b' X_j^V /s_j). Using the property that $E(b'X_{ij} - b' X_j^V) = 0$, we obtain (Maddala (1983)):

$$E(H_j) = E(\Phi(-b'X_{ij}/s_j))$$

$$\approx \Phi(-b' X_j^V/s_j) + ½ \bullet (b' X_j^V/s_j^3) \bullet \phi(-b' X_j^V/s_j) \bullet E(b'X_{ij} - b' X_j^V)^2 \quad (5)$$

where $\phi(.)$ is the normal density function and $E(b'X_{ij} - b' X_j^V)^2$ is the variance of the predicted household level consumption around the predicted mean consumption within each village. In words, the predicted level of poverty for villages outside the sample is a function of the mean level of consumption per adult and of its variance around that mean. Equation (5) can therefore provide a prediction of the incidence of poverty in communities outside the sample, using the estimates (b and s_j) of the parameters of the household consumption function and the community-level characteristics X_j^V of the villages outside that sample.

Note that the regression analysis is used to predict consumption levels for all households rather than whether or not the household is poor. The latter approach would be equivalent to estimating (1) directly. This is often referred to as a multivariate 'poverty profile' (Ravallion (1996)). The individual poverty indicator in (1) is binary, so one could use a probit (or alternative) model to construct a prediction model. As pointed out by Ravallion, a puzzling feature of this approach is that the estimation techniques used were typically developed for situations in which the observed data were dichotomous or truncated at zero. The standard way of writing the solution to this estimation problem is then to define a regression model in which a continuous latent (*unobserved*) variable is regressed on a set of observed explanatory variables (Maddala (1983)). A particular error structure (e.g. the normal distribution for the probit) is then assumed, allowing the parameters of inference to be estimated. These

village and the village-level variance of the prediction model.

[8] Hentschel et al. (1998) use this property to predict regional poverty from census data.

parameters can then be used for inference related to the explanatory variables and the *observed* limited dependent variable. If this procedure is used on a poverty indicator, such as the head count, then the latent variable is in fact an *observed* variable that was used to calculate the limited dependent variables. Since the latent variable is observed, limited dependent variable estimation of the poverty indicators is not necessary and will be less efficient, since some information actually available is not used in constructing the prediction model.

3. The Data

Data for this study were collected from a large number of sources and brought together at the level of the village according to the name of the village and its geographical coordinates that indicate its location. Some of the data, most importantly the census, cover all the villages in the country or the entire population; other data cover only a sample of villages and a sample of sample of households within each village. Table 1 lists the different sources of data collected as well as their coverage. Not all data could be used in the econometric analysis, however; some of the data did not cover all provinces, while other data, most notably the Agriculture Survey, did not contain the information that was necessary in order to incorporate the data in the GIS database.

INSERT TABLE 1

After collecting all the data, they were standardized and integrated within a common data set. At the conclusion of this stage of the work, the database contained more than 60 tables that include data on the geographical coordinates of all of the country's villages, towns, markets and public facilities; data on the entire road network, socio-economic and demographic data from a variety of surveys and the population census, and a large data set on the agro-climatic conditions in the country's main provinces. The data were organized as a Geographical Information System (GIS), namely a computer system that allows the analysis and display of geographic and non-geographic data.

As an illustration of the type of information that was extracted from the Geographical Information System for the Community study, Map 1 shows the location of water points and their proximity to the villages in the Department of Karangasso-Vigue. The points in the map that indicate the location of the villages are scaled according to the size of their population, thus showing the demand pressures on each water point. The map also contains information on the road infrastructure, including the quality of the roads, and on the hydrographic networks.

INSERT MAP 1 (water)

The application of these data for predicting poverty across communities in the country, had to be constrained in this study to a smaller data set since not all data were available for all villages at the time that the data were collected. In particular, Table 1 indicates that the data that were obtained from the Ministry of Water Management were limited to 25 provinces, or 5207 out of the country's 6821 villages. Data for the remaining 5 provinces were subsequently collected from the Ministry of Water Management but these data were collected in another survey and there were very few variables which were comparable between the two surveys. In some villages there were missing data also on other variables and, for this reason, the number of villages in the final prediction analysis had to be reduced to 3871, or 57 percent of the country's total number of villages. Tables 2 and 3 provide the details of the variables used in the final analysis, emphasizing the limited coverage of some of the variables The lack of sufficient data for all of the country's 6821 villages is, of course, a cause of concern. Some of the missing data are available in the archives of the different ministries and, in principle, could be retrieved. However, if a significant number of villages still do not have all the necessary data, targeting will have to be made at higher levels of geographical aggregation of the department or the provinces. Targeting at these levels would still have to use the predictions of the village-levels of poverty which were obtained in this study for all the villages outside the sample since it cannot make a direct use of the Household survey. The reason is that the sampling frame and the sample size do not provide an adequate representation of all departments and provinces.

Similarly to the LSMS surveys, the sampling for the Priority Survey used in the estimation of the consumption model was semi-stratified (INSD, 1996). The survey was designed to be representative at both the national and regional levels. First, the country is divided into seven regions. Five rural regions that represent five agro-climatic areas and two urban regions: one comprising Ouagadougou and Bobo-Dioulasso, the two main cities, and the other one comprising all the other remaining cities. 434 enumeration areas were selected from all the seven regions on the basis of socio-economic characteristics.

In each of the 434 enumeration areas, 20 households were randomly selected. In the econometric estimation, however, the sample size had to be reduced to 5618 households and 201 enumeration areas as a result of the missing data for 5 provinces and the incomplete data on certain variables in a few other villages.

INSERT TABLES 2 and 3

4. The empirical model

As discussed in section 2, the econometric analysis has two parts. First, a consumption model is specified for the households included in the Priority Survey. Secondly, this model is used to predict for poverty levels by community for all communities included in the GIS database. Table 2 presents the data that were used in the econometric analysis. The corresponding community-level variables that were used for the prediction are presented in Table 3. To estimate household consumption levels, we used a standard reduced-form framework in which income (measured in terms of household consumption) is regressed on household characteristics, including human and physical capital, as well as on community level characteristics.[9] Some community characteristics are specified at the village level, whereas others, primarily the agro-climatic conditions, are specified at the department or region levels.

The Priority Survey contains a limited but important set of variables that can be used to explain households' consumption levels. In this study, the household-level explanatory variables had to be selected so that they allow aggregation at the community level, and thus be used for the prediction. This limited the choice of explanatory variables to only those household-level variables for which the corresponding mean values at the community level were available for all communities in the country. As a consequence, several variables such as the level of education of the household's members (as opposed to the literacy of the household's head) — which are usually found to be significant in a consumption model — could not be included in the estimation. Furthermore, data on household assets and land holdings, which are also significant explanatory variables in most consumption models, were not available in the Burkina Faso PS. This reduced the

[9] Examples are Glewwe and Kanaan (1989), or Coulombe and McKay (1996). Glewwe (1991) has a useful discussion on the justification for including particular variables in this type of approach. We will return to the problems related to this specification below.

explanatory power of the model, and very likely created also a standard missing variable problem. Moreover, since the underlying data are cross-section, household heterogeneity — a common problem affecting any regression of welfare indicators — is also hard to address. Despite these reservations, we were able to collect data on most of the important explanatory variables and include them in the model (see table 2). They include demographic variables, variables that characterize the household's human capital such as literacy rates, and the household's physical capital such as livestock.

Table 2 also lists the variables at the village- and the department-level variables that were included in the econometric estimation. Department-level variables are primarily climatic data and department-level means of certain household variables (e.g., average area of cultivated land in the Department) which were obtained from the Ministry of Agriculture. In the analysis of the impact of climatic factors, we distinguished between the impact of the long-term climatic characteristics and the impact of temporary fluctuations by including among the explanatory variables the average level of rainfall during the past 15 years as well as the absolute value of the deviation of the past year's level of rainfall from the long-term average. The village level explanatory variables include also data on the distance to schools and health facilities, the quality of the access road, the quality of these facilities, and water supply in the community.

5. Estimating Poverty within the Sample

Table 4 provides some descriptive statistics on poverty and consumption per adult equivalent by agro-ecological zones in rural and urban areas of Burkina Faso. The poverty line in these estimations was equal to two-thirds of the country's mean level of consumption per adult equivalent. The table emphasizes the large difference in the incidence of poverty between urban and rural households and the much higher standard of living in the country's two main cities. In rural areas, the western region has relatively lower rates of poverty, whereas in the other regions, poverty rates are higher and the differences between these regions are rather small.

INSERT TABLE 4

In the econometric analysis, consumption per adult equivalent was regressed on the explanatory variables listed in table 2, according to the linear model in Equation (2). The model was estimated via the maximum likelihood method, in which the regression coefficients and the heteroscedastic errors were jointly estimated.[10] By allowing for heteroscedasticity by community, the community level information can be used for predicting the mean level of consumption per standard adult as well as the variations of consumption around this mean level. This estimate of the variance within a community may provide some information on the extent of inequality in the distribution of consumption within that community. The regression analysis was conducted separately for households in rural and urban areas.[11] The results for the consumption regression are reported in Table 5a, and the results for the errors regression are reported in Table 5b. In the regression results for both urban and rural areas, multiplicative heteroscedasticity cannot be rejected at 1 percent.[12]

Because maximum likelihood estimation was used to jointly determine the coefficients in the model and the heteroscedasticity structure, no simple R^2 can be reported. The first step OLS estimates of the model indicate, however, that the adjusted R^2 is quite low, with R^2 equal to 0.28 for the urban population and equal to 0.17 for the rural population. These low levels are primarily due to the restrictions on the choice of variables that were included in the model in order to assure that these variables are available for all communities outside the sample. When *all* the household- and community-level variables which were available in the PS were used in the estimation, the value of R^2 in the regression for rural households rose to 0.50. The low values of the adjusted R^2 with the more limited set of explanatory variables required us to make significant adjustments in the application of the results in the prediction model. These adjustments are discussed in the next section. The results show, however, that the variables included in the model are strongly jointly significant and that a substantial number of household- and community-level variables are highly significant. The following results stand out:

- The household-level variables that are most closely correlated with the level of consumption in both rural and urban areas are the literacy rates of the adult members of the household.[13]

---

[10] The regression was weighted with individual sampling weights derived from the original sampling frame used by the World Bank/INSD.
[11] Pooling tests convincingly rejected running one national regression.
[12] The Breusch-Pagan LM-test convincingly reject homoscedasticity (see Table 5a). The Glesjer (1965) test indicates that in both urban and rural areas, the null hypothesis of multiplicative heteroscedasticity cannot be rejected at the 1 percent level.
[13] Note that the variables describing the literacy of adults in the households also include the household

- The dependency rates, namely the number of children and elderly persons per adult in the household, do not seem to have a clear effect on consumption of rural households whereas for urban households, dependency rates, particularly the number of *elderly persons* per adult in the household, have a significant impact by reducing the level of consumption per adult equivalent.
- Livestock units per capita — the only proxy for the household's physical assets available in the Priority Survey — is found to be significantly and positively correlated with consumption in rural areas.
- The community-level variables in rural areas that characterize the agro-climatic conditions have a strong impact on consumption. In rural areas that have relatively high levels of *long-run average* rainfall, relatively normal *rain in the survey year*, and low rainfall *variability* over the rainy season, consumption per capita is typically higher.
- Interestingly, though, in the urban consumption function the agro-climatic variables that indicate the average level of rainfall and the homogeneity of the rainy season seem to have a *negative* effect on consumption. A possible explanation is that the consumption basket of these households include, in normal years, commodities that were not recorded in the PS survey (which includes only a small number of consumption items); the reduction in consumption during the normal years recorded in the survey is therefore spurious. Another possible explanation is that these variables are correlated with certain significant *missing* variables that have a negative impact on the consumption of urban households. The data we had at our disposal did not allow us, however, to make a further analysis of these effects, however, intriguing.
- In rural areas, the level of consumption in villages that are further away from schools is generally lower. No similar effect is revealed with respect to the distance to health facilities. A possible explanation is that in some regions, villages located further away from the health facility receive services from mobile health clinics.
- In both urban and rural areas, the *quality* of the services in the health facility — approximated by the variable that indicates the availability of safe drinking water in the facility — is significantly correlated with the level of per capita consumption in the surrounding villages and urban neighborhoods. Only about one-third of the health

---

head.

facilities in Burkina Faso have safe drinking water.

- In rural areas, the quality of the infrastructure, indicated by the availability of safe drinking water in the village — measured by the number of functioning pumps — and the quality of the access road to the village, has a significant and positive impact on consumption: in villages that have access to an all-weather access road, mean consumption is nearly 10 percent higher than in villages that do not. The greater opportunities to trade, rather than produce for own-consumption, and the better alternatives for non-agricultural work that the access to an all-weather road provides are the main reasons for this effect.

INSERT TABLE 5a

- The coefficients that determine the pattern of the village level error terms indicate that in villages with a relatively high proportion of literate heads of households, the distribution of per capita consumption is less equal than in villages with a low proportion of literate heads of households possibly -- because in a village with more literate adults, the income differences between households with less educated heads and households with more educated heads are relatively larger.
- In villages with relatively high average levels of rainfall there are larger differences between households in their levels of per capita consumption, possibly because in these villages some households are better equipped and more capable to take advantage of the better conditions for agriculture.
- Villages with higher average land holdings per household have a larger variability of per capita consumption.

INSERT TABLE 5b

Several of the above interpretations of the results suggest possible *causal* relationships between the explanatory variables and the dependent variable. However, these interpretations are intended primarily as a background for a more thorough evaluation of the possible policy implications of the results, and they are made under the usual caveat of possible endogeneity of the community level variables, which means that correlation need not be an indication of causality. Thus, for example, government policy of locating relatively more public education facilities in the relatively poor villages as part of an anti-poverty program, will lead to high *negative* correlation between the average level of per capita consumption and the proximity of the village to school (Rosenzweig and Wolpin

(1986)). The availability of a large number of water pumps in a village, to take another example, need not be the *cause* of a relatively high standard of living of the households in that village, but rather the *effect* of the larger demand for safe drinking water of the more affluent villagers.

In the present analysis, the objective of the regression estimates is, however, to construct a prediction model that can identify the very poor and the least poor villages. The quality of these predictions depends only on the degree of *correlation* between the explanatory and the dependent variables irrespective of whether or not this high correlation indicates causality. If, for example, health facilities are systematically placed in the poorer villages, then the variable that indicates the distance from the village to the health facility can be useful for predicting the standard of living of the households in these villages[14]. Nevertheless, the possibility of endogeneity requires special care in the interpretation of the results for policy purposes[15]. While the significance and size of the coefficients are suggestive, more additional work would be needed for designing appropriate poverty reduction policies.

6. Predicting the Geographical Distribution of Poverty in Burkina Faso

The next step is to apply the regression results that were obtained for a sample of communities to the data available in the GIS database for *all* communities in Burkina Faso in order to predict the distribution of poverty across all communities. as noted earlier, in this analysis, we had to focus on the 3871 out of a total of more than 6,000 villages for which all the necessary information was available to us. The Headcount index of a community is calculated by means of Equation (5), using the estimates of the parameters that were obtained in the regression analysis. To apply this equation, it was necessary to provide estimates for the level of mean consumption per adult in the community and for the variance of consumption. For the term $E(b'X_{ij} - b' X_j^V)^2$ in Equation (5), we used the average value per region (rather than per community) in the survey data, while s_j was obtained using the coefficients given in table 5b. Mean

[14] Note that this requires further that the same program placing rule is used outside the sample as inside the sample. Since the sample is nationally representative, this may be an appropriate assumption.

[15] There are other sources of endogeneity possible. For example, our approach assumed that location is not a choice variable. Migration is therefore not explicitly considered, requiring further care in the interpretation of the results.

consumption per standard adult in the communities outside the sample was predicted from the mean values of the explanatory variables for each of these communities, using the coefficients given in table 5a.

Before applying these predictions for all the communities outside the sample, we assessed the quality of the predictions by comparing them with the direct estimates of poverty in the sample of 201 communities which were included in the PS. Toward that end, we calculated the correlation coefficients of the predicted and calculated poverty levels for these villages. The value of the Pearson-correlation coefficient was 0.51 and it was strongly statistically significant. For policy decisions, the more relevant criterion is the relation between the *order* of villages on the poverty scale according to the incidence of poverty that is determined by the predictions, and the order determined by the direct estimates from the PS survey. To test this aspect of the prediction model, we calculated the Spearman rank correlation coefficient between the order established by the direct estimates of poverty and the order established by the model's predictions. That correlation coefficient was also strongly significant at 0.43. Table 6 provides another illustration of the quality of the predictions by comparing the estimated and the predicted values of the Headcount measure of poverty for selected communities in three provinces. Although the predicted values of the headcount measures per community often fall outside the confidence interval of the calculated levels of poverty, the rank order of communities from the richest to the poorest in each province is quite similar.

INSERT TABLE 6

Despite these results of the tests, the low levels of R^2 in the regression analysis of Equation 5, and the low quality of the data prevented us from making a direct use of these predictions. Moreover, these predictions rely on the assumption of normality of the error term. One common test for normality is the Jarque-Bera test; our estimate of the Jarque-Bera statistic was 11.8, and the normality hypothesis therefore had to be rejected.[16] As a result, we did not use the prediction in order to establish a *complete* order of the communities on the poverty scale. Instead, we divided the 3871 villages and

[16] This statistic is distributed a Chi-square with two degrees of freedom and the normality hypothesis had to be rejected at 0.997 probability. One should note, thoughm that the Jarque-Bera test is not robust to the presence of heteroskedasticity, which could not be rejected by the Breusch-Pegan LM test and the Glesjer test. We are not aware of a test of normality in the presence of heteroskedasticity, but the high value of the Jarque-Bera statistic suggest that it is highly probable that the residuals are not distributed normally.

urban communities for which predictions were made into four categories of poverty, ranging from the poorest to the least poor, according to the predicted levels of poverty. Despite the errors in these predictions, our results suggest that the majority of the villages in the lowest category of the poorest villages are likely to have a higher incidence of poverty than the majority of the villages in the highest category of the least poor villages. The villages in the ‘poorest’ category are therefore candidates for targeted poverty alleviation programs and the villages in the ‘least poor’ category are candidates for cost-recovery program. Given the data limitations in Burkina Faso, effective targeting would have to focus only on these two *extreme* categories in order to reduce the leakage as much as possible and keep the budget constraint. Further improvements in targeting in the present circumstances when the available data are rather limited and their quality quite low can be achieved by dividing the villages into a larger number of categories and focusing on the villages in the two extreme categories; future research aimed at improving targeting would have to focus, however, on efforts to improve the quality of the data as well as generate additional series of geo-referenced data.

Table 7 presents the geographical distribution of the rural and urban communities across these categories of well being in the different provinces of Burkina Faso. The Table was constructed by dividing the villages into the four categories — ranging from the ‘poorest’ to the ‘least poor’ — using the predicted values of the poverty incidence, and allocating the entire population in each of the villages to the corresponding category. The categories were determined so that, in the country at large, the population in each category represents 25 percent of the country’s total population. The distribution of the population in the different provinces is significantly different, however. For example, 41 percent of the population in the province of Boulkiemde lives in villages that were classified in the category of the ‘poorest’ villages, and only 7 percent of the population in this province live in villages that were classified in the ‘least poor’ category.

Consider, as an illustration, an anti-poverty program targeted on five of the provinces in which at least 40 percent of the population reside in villages that were classified in the ‘poorest’ category. Under this criterion, 21 percent of the country’s total population will be included in the target provinces. Only 3 percent of the population in the four target provinces (which account to only 0.6 percent of the country’s total population) live, however, in villages that were classified in the *‘least poor’* category — suggesting that leakage is likely to be quite small. Nearly 43 percent of the population in the four target provinces live in villages that were classified in the ‘poorest’ category and they account for 36 percent of the country’s total population that live in the ‘poorest’ villages. At the other extreme, a cost-recovery program that is targeted on 7 of the provinces in which more than 40 percent of the population live in villages that were classified in the ‘least-poor’ category, will cover 26 percent of the country’s total population but only 13 percent of the country’s population that live in the ‘poorest’ villages.

Targeting anti-poverty programs at the province level is likely, however, to be less effective than targeting at the village level. The reason is that province targeting is bound to include villages of the higher categories in which the incidence of poverty is likely to be lower than that in the villages of the lowest category. Under a given budget constraint, a targeted program at the village level that focuses only on villages of the ‘poorest’ category is therefore likely to cover a larger share of the country’s poor population and entail less leakage than a program targeted at the province level — despite the prediction errors in the classification of the villages into these categories.

In urban areas, most of the urban communities were classified in the ‘least poor’ category. This result is largely due to the much higher standard of living and much lower incidence of poverty in urban areas. There are several other, more technical, reasons, however, for this classification:

- In urban areas the distinction between ‘poor’ and ‘non-poor’ *communities* is less clear than in rural areas — first, because, in many developing countries, it is not uncommon to have poor households that reside in relatively affluent urban communities and vice versa, and second, because the *communities* — as defined in the household surveys — are, in fact, *enumeration areas* that have been determined by the local authorities for administrative purposes and their borders are often quite arbitrary. Whereas in rural areas the enumeration areas are generally limited to one or two neighboring villages which tend to have similar living standards, in urban areas, where the distance between neighborhoods is small, enumeration areas often include communities with largely different living standards. In our study, we had access to community level data in urban areas only in the Household survey.
- In all other data sources, the towns, including Ougadougou and Bobo-Dioulasso, were considered as single points in the GIS data-set. In the econometric analysis, all

the enumeration areas from each of the large towns had the same community characteristics and thus had to be considered as a single 'entity.'

INSERT TABLE 7

The main advantages of community targeting are demonstrated in Map 3 which shows all the villages in the Province of Sanguie — divided into the four categories of well-being. The map shows that in this area, most of the 'poorest' villages are located further away from the urban centers and they are not connected to an all-weather road. The map also highlights the fact that targeting an anti-poverty program on the entire population of the province of Sanguie is bound to include many non-poor villages, whereas excluding this province from the program will leave out a considerable number of poor villages.

INSERT MAP 2 (poverty)

7. Simulating the impact of a village-level targeting scheme

To evaluate the effectiveness of community targeting of anti-poverty programs, we conducted a simple simulation experiment. For this experiment, we use the data on consumption in the 201 communities for which we have complete information from the Priority Survey. The simulation design follows closely the framework of Baker and Grosh (1994). It is assumed that the government has a given budget for income transfers to the target population. The effects of the income transfers on poverty are evaluated using the *actual* household consumption data. The selection of the villages for targeting is determined, however, by the *predicted* levels of the village poverty that have been determined by the model. The simulations can thus evaluate how effective were these predictions in *identifying* the poor, by estimating the leakage, i.e. the number of non-poor included in the scheme and the undercoverage, i.e. the proportion of poor not reached by the program. The estimates are made for the households included in the survey but we use the individual sampling weights in order to measure the impact on the total population at the regional and national levels. The reliability of the regional and national results is therefore affected by the sampling errors that are due to the sampling frame of the PS survey.

To evaluate the program, its outcomes are compared with a untargeted uniform transfer scheme, in which all the individuals in the country receive a transfer. We also consider two other targeted programs: a village-level targeted program that uses the actual poverty levels to identify the poor villages included in the program and a 'perfect' targeting program that uses the actual household consumption data to identify the poor included in the program. The three targeted programs are designed to include 30 percent of the population.[17] To achieve this, we first set the poverty line so that 30 percent of the total population in the country is identified as poor.

Selection into the three targeted programs is as follows. For the program of village-level targeting, based on the predictions of the village levels of poverty, all the villages in the sample were first ranked according to their levels of poverty. The villages are selected for targeting starting from the poorest village until (at least) 30 percent of the population is being included in the program. For the program of village-level targeting, based on the actual village levels of poverty, all the villages were also ranked according to their poverty levels and the poorest villages are included until (at least) 30 percent of the population is being included. For the program of household-level targeting, based on the actual consumption per adult equivalent data, households were ranked according their consumption levels. Households are included for targeting starting from the poorest until 30 percent of the population is being included in this program.

Table 8 presents the results for the undercoverage and the leakage implied by the different targeting programs[18]. First, consider a targeted program using actual village level poverty levels. The simulation results for this program suggest that 44 percent of the poor would not be covered whereas fully 56 percent of the poor would be covered. With the village-level targeting that is using the *predicted* poverty-levels per community, the undercoverage would rise to 56 percent as an effect of the prediction errors but 44 percent of the poor population would be covered nonetheless. As an indicator of the accuracy of the predictions, this implies that about 79 percent of the poor that could be reached via targeting using *actual* village-level poverty data could be reached via targeting using the *predicted* poverty-levels per community.

INSERT TABLE 8

By design, untargeted transfers result in zero undercoverage but leakage is high. Individual leakage is defined as the number of non-poor covered by a particular program

[17] The simulation uses a lower poverty line than in the previous section to focus on attempts to target a relatively small part of the population. Given that poverty using 2/3 of the mean as the poverty line was estimated 58%, this would have suggested a transfer program that attempted to include two-thirds of the population, making the issue of undercoverage and leakage less interesting to study. Note that the total population considered in this simulation is the households for which we have the complete set of variables in the prediction model.

[18] Further details, including on the inter-regional distribution of leakage and undercoverage is given in a longer, working paper version of this paper, Bigman et al. (1999).

divided by the total number of people reached under this targeting rule. For the untargeted program this is (by design) 70 percent, since 30 percent of the population are considered poor in these simulations. When using the village-level poverty estimates based on the PS data to establish criteria for targeting, the leakage is reduced to 44 percent. Note also that the distribution of leakage across the (rural) regions is very similar. When using the predictions on village-level poverty to select the communities included in the scheme, the leakage increases to 56 percent. Still, this is far lower than the leakage implied by undifferentiated transfers[19].

8. Conclusion

Geographical targeting of anti-poverty programs can provide an effective strategy of reaching the poor and keeping the costs of anti-poverty programs in check in countries where the information on individual households is incomplete or unavailable and a practical individual or household targeting is therefore not possible. By identifying the geographical areas in which the poor concentrate, these programs can reduce the leakage to the non-poor so that a larger share of the poor population can be reached with a given budget and a larger share of this budget can reach these poor. However, in most countries where geographical targeting is applied, the target areas are the region, the state or the entire rural area. Although targeting even at these levels can offer considerable savings compared with a non-targeted program that cover the entire country, they necessarily involve substantial leakage to non-poor households that reside in the target areas. Narrow targeting at the level of the community or the administrative department can offer an more effective alternative of reaching the poor by reducing the leakage and lowering costs. Narrow targeting can be effective for two main reasons: first, in most developing countries, particularly in Sub-Saharan Africa, poverty tends to be concentrated in villages and certain parts of towns. Second, geographical targeting

[19]In the Bigman et al. (1999), we also introduce actual income transfers and evaluate the poverty impact of the transfers for a given budget. This budget is equally divided amongst all the individuals included in each program. Using a budget of just over half the actual poverty gap, undifferentiated transfers reduce the headcount measure of poverty by a relatively high 22 percent, while village-level targeting using the actual poverty levels reduced that measure of poverty by 33 percent. Using predicted poverty levels, one can reduce poverty by about 27 percent. These are only modest gains, but note that giving *all households* living in communities included in a program *the same money transfer* is by no means optimal when minimizing poverty. Our scheme is aimed at *identifying* the poor communities and it attempts to minimize leakage and undercoverage.

requires relatively low costs to administer the programs and, by relying primarily on the local authorities, has the potential of securing that the larger portion of the benefits will reach the target population.

This paper presents a methodology of using data from a wide variety of different sources in order to establish criteria for targeting poverty-reduction programs at the levels of the village, the urban community, or the local administrative department. This methodology is demonstrated in the paper for Burkina Faso. It consists of collecting data from a large number of sources, bringing them together at the village level, and arranging them as a Geographic Information System. Data on the population were collected from several of socio-economic and demographic surveys as well as the population census; data on the road infrastructure, on public facilities, on the location of central towns and markets were collected from several government ministries and public administrations; agro-climatic data were collected from local and international research institutes. An econometric analysis was then conducted with the data of the household survey to identify the variables that best explain the households' consumption levels. The explanatory variables in this analysis included important characteristics of the community — such as the distance to the urban center and the public facilities, the quality of the access road and the agro-climatic conditions — together with key characteristics of the households in that community — such as literacy rates or dependency ratios. The explanatory variables at the household level were selected so that their mean values per community were available for the majority of the communities in Burkina Faso outside the household survey. This made it possible to use the model that has been estimated in the regression analysis with the data of the PS in order to predict the incidence of poverty in all the villages outside the PS sample and thereby identify the spatial distribution of poverty at the community level.

In the present analysis for Burkina Faso, constraints on the availability and quality of the data led to considerable prediction errors and prevent us from using the complete ordering of the villages according to the incidence of poverty as has been predicted by the econometric analysis. We used a simple method to reduce the impact of these errors by dividing the villages into several categories and focusing only on the categories of the poorest and the least poor villages. Indeed, practical considerations in the application of anti-poverty programs and tight budget constraints are likely to reduce the need for the complete ordering. Instead, poverty alleviation programs are likely to focus on villages at the lower end of the distribution, and cost-recovery programs are likely to focus on the villages at the higher end. Nevertheless, the limited availability of geo-referenced data and the low quality of the data currently available reduced the predictive power of our econometric analysis, and further work would be necessary in order to augment and improve the stock of relevant data.

Targeting poverty alleviation or cost recovery program at lower-level administrative areas of the village or the department have other advantages as well: First, budget constraints are likely to restrict programs that are targeted on larger geographical areas of regions or states, and, as a result, the errors of inclusion and of exclusion with the latter programs are likely to be quite high. Targeting on smaller geographical areas can reach, with the same budget constraints, many more of the country's poor. Second, lower level targeting is likely to include villages and districts in all regions or states and thus be less divisive and contentious on ethnic, social or political grounds. Third, whereas the differences in the incidence of poverty between regions are primarily due to differences in their agro-climatic conditions, differences in the incidence of poverty between villages within the same region often reflect past biases in policies that led to differences in the quality of their access road or their public services; targeting future policies in light of these criteria can remedy these past biases.

Bibliography

Alderman H. (1987), "Allocation of goods through non-price mechanisms: Evidence on distribution by willingness to wait", *Journal of Development Economics* 25(1), 105-124.

Baker, J. and M. Grosh, (1994),"Measuring the effects of geographic targeting on poverty reduction," LSMS Working Paper 99, World Bank.

Besley, T. and R. Kanbur, (1991), "The principles of targeting", in Balasubramanyam V. and Lall, . S. (eds.), *Current issues in Development Economics*, 69-90.

Bigman, D., S.Dercon, D.Guillaume and M.Lambotte (1999), "Community Targeting for Poverty Reduction in Burkina Faso", *CES Discussion Paper Series, DPS 99.10*, Center for Economic Studies, Leuven.

Coulombe, H. and A. McKay, (1996), "Modeling determinants of poverty in *World Development* 24, 1015-1031.

Deaton, A., (1997), *The Analysis of Household Surveys: a Microeconometric Approach to Development Policy*, Washington D.C., The World Bank and Johns Hopkins

University Press.

Glesjer, H. (1965), "A New Test for Heteroscedasticity", *Journal of the American Statistical Association*, 60, 539-547.

Glewwe, P., (1991), "Investigating the determinants of household welfare in Cote *Journal of Development Economics* 35, 307-337.

Glewwe, P. and O. Kanaan (1989), "Targeting Assistance to the Poor Using Household Survey Data", Policy, Planning and Research Working Papers, WPS 225, World Bank

Greene, W.H. (1993), *Econometric Analysis*, 2nd edition, New York: Macmillan

Hentschel, J., Lanjouw, J., Lanjouw, P. and J. Poggi, (1998), "Combining survey data with census data to construct spatially disaggregated poverty maps: a case study of Ecuador", mimeo, World Bank.

INSD, (1996), "Le Profil de pauvrete au Burkina Faso", Ouagadougou.

Maddala, G. S., (1983), *Limited dependent and qualitative variables in econometrics*, Cambridge University Press.

Ravallion, M., (1993), "Poverty alleviation through regional targeting: a case study for Indonesia", in K. Hoff, A. Braverman and J. Stiglitz, *The economics of rural organization,* Oxford University Press.

Ravallion, M. (1996), "Issues in Measuring and Modelling Poverty", *Economic Journal*, 106, September, 1328-1343.

Rosenzweig, M. and K.Wolpin (1986), "Evaluating the effects of optimally distributed *American Economic Review*, 76 (3), 470-482.

van de Walle, D., (1995), "Public Spending and the Poor: Theory and Evidence: Introduction", in van de Walle and Nead (eds.), *Public Spending and the Poor : Theory and evidence*, Baltimore : Johns Hopkins University Press, 1-5.

Table 1: Data sources

Level of aggregation	*Data Source*	*Acronym*	*Coverage*
Household	Priority Survey (1994): provides data on income and expenditure for 8642 households	PS	survey sample (473 villages)
Village	Priority Survey (1994): community component of the PS which covers infrastructure and communal services	PS	survey sample (473 villages)
Village	National census (1985): demographic data	NC	national
Village	Ministry of Water Management and Infrastructure (1995): data on health and water infrastructure, distances to infrastructure, public administration and social groupings	DGH	25 out of 30 provinces
Village	Ministry of Education (1995): data on primary school infrastructure and teacher/pupil ratios.	EDU	national
Department	Ministry of Agriculture (1993): data on various indicators ranging from average literacy rates to vegetation indices	ENSA	national
Department	Directorate of Meteorology (1961-1995): data on temperature (31 locations), evapo-transpiration (15 loc.) and rainfalls (160 loc.).	METEO	national
Province	Ministry of Agriculture (1993): data on cattle per households	ENSA	national

Table 2. Descriptive statistics on variables used in the estimation

*Aggregation Level **	*Variable*	URBAN			RURAL			*Data source ***
		Mean°	*Standard Error°*	*Number observ.**	*Mean°*	*Standard Error°*	*Number observ.**	
Household	children 0-6 per adult (15-50 years) in household	0.530	0.495	2671	0.779	0.598	5508	PS
Household	children 7-14 per adult in household	0.618	0.590	2671	0.748	0.640	5508	PS
Household	elderly (50+) per adult in household	0.183	0.343	2671	0.313	0.426	5508	PS
Household	literate head in household	0.477	0.499	2736	0.134	0.341	5906	PS
Household	% male adults literate in household	0.562	0.422	2736	0.177	0.313	5906	PS
Household	% female adults literate in household	0.373	0.397	2736	0.053	0.174	5906	PS
Household	livestock units per capita	0.123	0.909	2736	0.442	0.943	5906	PS
Village	distance to nearest rural primary school				2.29	5.64	4412	DGH
Village	teachers per child 7-14 years	0.014	0.002	2736	0.005	0.006	5760	EDU
Village	distance to nearest health facility				4.79	7.77	4434	DGH
Village	whether nearest facility has safe water	0.82	0.39	2416	0.034	0.18	4434	DGH
Village	number of pumps per rural community				7.35	10.64	5241	DGH
Village	existence of an all-weather road				0.57	0.50	4434	DGH
Department	cultivated area in department per capita	0.211	0.221	2736	0.507	0.301	5760	FEWS
Department	average rainfall 80-94	65.80	10.07	2736	62.50	14.84	5760	METEO
Department	94 absolute value of deviation of rainfall from average	19.45	14.49	2736	22.58	12.96	5760	METEO
Department	average length rainy season 82-92	9.52	1.34	2736	9.53	2.00	5760	FEWS
Department	average vegetation index 82-92	0.114	0.034	2736	0.136	0.051	5760	FEWS
Department	homogeneity rainy season 82-92	0.162	0.019	2736	0.161	0.036	5760	FEWS

For community level variables, the same value is assumed for all the households of the community. ** acronyms are given in Table 1 on data sources.
° weighted using sampling weights.

Table 3. Descriptive statistics on variables used in the prediction

*Aggregation Level **	*Variable*	URBAN			RURAL			*Data source ***
		Mean°	*Standard Error°*	*Number observ.**	*Mean°*	*Standard Error°*	*Number observ.**	
Village	children 0-6 per adult (15-50 years) in household	0.656	0.110	300	0.645	0.227	6818	NC
Village	children 7-14 per adult in household	0.593	0.120	300	0.563	0.280	6818	NC
Village	elderly (50+) per adult in household	0.320	0.076	300	0.348	0.351	6818	NC
Province	literate head in household	0.450	0.181	191	0.113	0.075	6711	PS
Province	% male adults literate in household	0.522	0.147	191	0.141	0.079	6711	PS
Province	% female adults literate in household	0.323	0.149	191	0.044	0.034	6711	PS
Province	livestock units per capita	0.147	0.090	191	0.492	0.263	6711	AGRI
Village	distance to nearest rural primary school				4.39	5.04	4556	DGH
Village	teachers per child 7-14 years	0.023	0.032	295	0.003	0.011	4753	EDU
Village	Distance to nearest health facility				6.79	7.46	4393	DGH
Village	whether nearest facility has safe water	0.15	0.35	219	0.005	0.073	4390	DGH
Village	Number of pumps per rural community				2.350	2.765	5425	DGH
Village	Existence of an all-weather road				0.43	0.50	4618	DGH
Department	Cultivated area in department per capita	0.669	0.605	300	0.751	0.717	6821	FEWS
Department	Average rainfall 80-94	65.52	15.34	300	69.16	16.34	6821	METEO
Department	94 absolute value of deviation of rainfall from average	18.90	11.54	300	18.61	13.95	6520	METEO
Department	Average length rainy season 82-92	10.19	2.31	300	10.77	2.44	6520	FEWS
Department	Average vegetation index 82-92	0.126	0.054	300	0.121	0.051	6821	FEWS
Department	Homogeneity rainy season 82-92	0.152	0.038	300	0.153	0.036	6821	FEWS

* For department or province level variables, the same value is assumed for all the households of the community. ** acronyms are given in Table 1 on data sources.
° weighted using total population relative to village population.

Table 4. Poverty and consumption in Burkina Faso: estimates of the Priority Survey (1994)

Level of aggregation	*Consumption Per adult Equivalent*[a]	*Headcount*[b]
Ouest	7573	0.56
Sud/Sud-Est	5699	0.67
Centre-Nord	4952	0.74
Centre-Sud	5240	0.75
Nord	6122	0.64
Other Urban	12173	0.39
Bobo/Ouaga	20768	0.14
Whole country	*8766*	*0.58*

[a]in F. CFA per month

[b]Poverty line is set at 2/3 of mean consumption.

Table 5a. Regression results - dependent variable is log(consumption per standard adult)

	Rural		*Urban*	
Variable	coeff.	t-value	Coeff.	t-value
Constant	7.71	52.48**	10.82	21.99**
Children 0-6 per adult (15-50 years) in household	0.02	1.55	0.01	0.40
Children 7-14 per adult in household	-0.03	-1.67+	-0.04	-1.60
Elderly (50+) per adult in household	0.03	1.24	-0.13	-2.91**
Literate head in household	0.18	3.66**	0.33	7.63**
% male adults literate in household	0.13	2.48*	0.16	3.18**
% female adults literate in household	0.55	8.41**	0.42	10.11**
Livestock units per capita(/10)	0.93	11.06**	0.31	1.49
Distance to nearest rural primary school(/100)	-0.48	-2.80**		
Teachers per child 7-14 years (*10)	0.21	1.07	-0.39	-0.22
Distance to nearest health facility(*100)	0.18	1.58		
Whether nearest facility has safe water	0.14	2.92**	0.92	5.24**
Number of pumps per rural community(/100)	0.34	3.27**		
All-weather road?	0.10	4.83**		
Cultivated area in department per capita	0.01	0.32	1.66	5.29**
Average rainfall 1980-94(/100)	0.53	3.39**	-1.64	-2.28*
1994 abs value of deviation from average(/100)	-0.22	-2.78**	-0.44	-2.03*
Average length rainy season 1982-92	-0.01	-0.69	-0.06	-0.98
Avg variab. Vegetation index 1982-92	-0.54	-1.81+	8.17	3.24**
Homogeneity rainy season 1982-92	2.50	8.38**	-12.26	-3.63**
F-joint significance regression	F[19,4107]	=34.58**	F[15,2346]	=53.70**
Number of valid observations	4119		2362	

**=significant at 1%; *=significant at 5%; +=significant at 10%

Table 5b. Regression Results: Estimated variance with Multiplicative Heteroscedasticity

	Rural		***Urban***	
Variable	coeff.	t-value	Coeff.	t-value
Constant	0.12	5.03**	0.03	1.75+
Children 0-6 per adult (15-50 years) in household. (community mean)	0.40	2.85**	-0.86	-3.68**
Children 7-14 per adult in household (community mean)	0.79	6.27**	1.12	5.08**
Elderly (50+) per adult in household (community mean)	-0.29	-1.64+	0.47	1.53
Literate head in household (community mean)	0.49	3.86**	0.09	0.91
% male adults literate in household (community mean)	-0.26	-1.96*	0.12	1.12
% female adults literate in household (community mean)	0.11	0.90	0.05	0.70
Livestock units per capita (community mean)	-0.05	-0.94	0.00	0.04
Distance to nearest rural primary school	0.00	0.07		
Teachers per child 7-14 years (*100)	0.20	4.70**	2.00	4.95**
Distance to nearest health facility	0.00	-0.97		
Whether nearest facility has safe water	-0.38	-3.28**	-1.47	-3.63**
Number of pumps per rural community	0.01	2.16*		
All-weather road?	0.20	4.03**		
Cultivated area in department per capita	0.33	3.05**	-1.03	-1.33
Average rainfall 1980-94	0.02	4.20**	0.02	1.38
1994 absolute value of deviation from average	0.00	1.14	0.00	-0.38
Average length rainy season 1982-92	0.09	2.78**	0.00	-0.01
Avg variab. Vegetation index 1982-92(*10)	0.31	4.36**	-3.19	-5.56**
Homogeneity rainy season 1982-92(*10)	-0.16	-2.25*	4.35	5.65**
Breusch-Pagan LM heteroscedasticity	=603.35**	(19 d.f.)	=158.33	(15 d.f.)
Glesjer-test multiplicative heteroscedasticity	F[19,4107]	=3.66**	F[15,2328]	=3.66**

**=significant at 1%; *=significant at 5%; +=significant at 10%

Note: A positive coefficient of the explanatory variable indicates that this variable has the effect of raising the variance; a negative coefficient indicates that this variable has the effect of lowering the variance.

Table 6. Comparison of the Model's Predictions of the Headcount Measure of Poverty and the Direct Estimates for Villages in the Sample for three Provinces

Province	***Village ID #***	***Within Sample Estimates****	**Outside Sample Predictions**
KOSSI	4426	0.28 (0.01)	0.24
	3786	0.33 (0.01)	0.67
	512	0.54 (0.01)	0.64
	2936	0.54 (0.04)	0.65
	5266	0.57 (0.02)	0.56
	5117	0.64 (0.02)	0.69
	1626	0.68 (0.02)	0.70
	1556	0.69 (0.01)	0.61
	1290	0.78 (0.01)	0.70
	250	0.80 (0.01)	0.57
KOURITENGA	744	0.64 (0.03)	0.66
	6233	0.72 (0.03)	0.66
	657	0.75 (0.01)	0.57
	1627	0.80 (0.03)	0.74
	2943	0.83 (0.01)	0.65
	3213	1.00 (0.00)	0.64
	3828	1.00 (0.00)	0.76
MOUHOUN	1278	0.23 (0.01)	0.32
	790	0.36 (0.02)	0.52
	6753	0.48 (0.03)	0.57
	4982	0.48 (0.02)	0.56
	740	0.50 (0.02)	0.52
	5149	0.57 (0.02)	0.52
	5635	0.72 (0.01)	0.60
	6674	0.72 (0.01)	0.65

*standard errors in brackets (Deaton (1997), p.47). The figures are weighted by household size.

Table 7. The Distribution of the Population in the Provinces of Burkina Faso into Four Poverty Categories According to the Classification of their Communities *

Province	*Poorest*	*Lower middle*	*Upper middle*	*Least poor*	*Total*		*% total population*
BAM	0%	10%	36%	54%	100%		0,90%
BAZEGA	13%	36%	35%	16%	100%		4,85%
BOULGOU	37%	22%	32%	9%	100%		7,04%
BOULKIEMDE	41%	36%	15%	7%	100%		5,99%
GANZOURGOU	47%	33%	19%	2%	100%		2,74%
GNAGNA	24%	13%	21%	41%	100%		3,73%
GOURMA	34%	17%	23%	25%	100%		5,55%
KOSSI	26%	30%	29%	14%	100%		6,22%
KOURITENGA	7%	25%	43%	25%	100%		3,11%
MOUHOUN	41%	39%	18%	2%	100%		6,06%
NAHOURI	23%	29%	40%	8%	100%		1,71%
NAMENTENGA	4%	25%	18%	52%	100%		2,16%
OUBRITENGA	9%	24%	40%	27%	100%		4,73%
OUDALAN	24%	36%	19%	21%	100%		1,69%
PASSORE	11%	2%	18%	68%	100%		3,77%
SANGUIE	31%	23%	25%	21%	100%		4,64%
SANMATENGA	17%	11%	25%	46%	100%		7,64%
SENO	26%	29%	19%	26%	100%		4,70%
SISSILI	41%	43%	16%	0%	100%		3,69%
SOUM	3%	14%	21%	61%	100%		2,54%
SOUROU	12%	12%	12%	64%	100%		5,31%
TAPOA	60%	17%	22%	1%	100%		2,12%
YATENGA	24%	24%	31%	21%	100%		6,72%
ZOUNDWEOGO	3%	59%	30%	8%	100%		2,40%

* Poverty line is set at 2/3 of mean consumption.

Table 8. Undercoverage and Leakage under Various Targeting Schemes[1]

	Undercoverage[2]	Individual leakage[3]
Untargeted transfers to the entire population	0.0%	70.0%
Village targeting using village level prediction	56.0%	56.1%
Village targeting using survey data only	43.9%	44.2%

[1]The schemes are designed to cover at least 30% of the population.
[2]Undercoverage gives the number of poor that are not included in the scheme divided by the number of poor that would be reached under perfect targeting.
[3]Leakage is the number non-poor covered by a particular targeting rule divided by the total number of people reached under this targeting scheme.